



JP8248985

Biblio

Page 1

esp@cenet**DEVICE AND METHOD FOR VOICE RECOGNITION**

Patent Number: JP8248985
Publication date: 1996-09-27
Inventor(s): NAKAGAWA SEIICHI
Applicant(s):: RICOH CO LTD
Requested Patent: ☐ JP8248985
Application Number: JP19950052165 19950313
Priority Number(s):
IPC Classification: G10L3/00 ; G10L3/00 ; G10L3/00
EC Classification:
Equivalents:

Abstract

PURPOSE: To reduce the amount of overall compute and to increase the processing speed of voice recognition by computing the output probability of a hidden Markov model(HMM) by a probability computing section based on the output probability read out of a probability storage section.

CONSTITUTION: A code book 8 which stores specimen vectors that are generated by time dividing specimen voices and a probability storage section 9, in which output probability for every state transition is stored corresponding to the specimen vectors, are beforehand formed. Then, inputted voices are time divided by a vector computing section 4, input vectors are successively generated and specimen vectors which are close to the generated input vectors are selected from the book 8 by a code selecting section 5. Plural output probabilities corresponding to the selected specimen vectors are read out of a probability storage section 9 by a probability computing section 6. The output probability of the HMM is computed by the section 6 based on the read output probabilities. Accordingly, during the computation of the HMM output probabilities, the amount of overall compute is reduced because the output probabilities are not computed from input vectors.

Data supplied from the esp@cenet database - I2

BEST AVAILABLE COPY

E 6055

(19) 日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平8-248985

(43) 公開日 平成8年(1996)9月27日

(51) Int. Cl. ⁴	級別記号	庁内整理番号	PI	技術表示箇所
G10L 3/00	535		G10L 3/00	535
	521			521C
	531			531G

審査請求 未請求 請求項の数 6 O L (全 9 頁)

(21) 出願番号 特願7-52165

(22) 出願日 平成7年(1995)3月13日

(71) 出願人 000006747

株式会社リコー

東京都大田区中馬込1丁目3番6号

(72) 発明者 中川 聖一

愛知県豊橋市王ヶ崎町字上原1-3 合同館

舎1-201

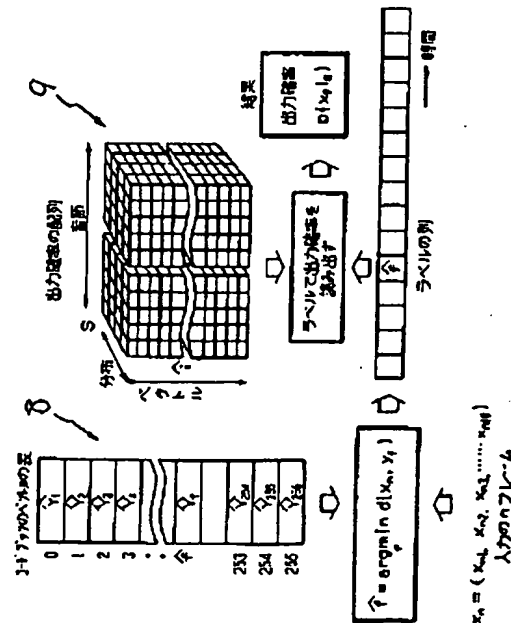
(74) 代理人 弁理士 柏木 明 (外1名)

(54) 【発明の名称】 音声認識装置及び方法

(57) 【要約】

【目的】 音声認識の計算量を削減する。

【構成】 標本音声の時分割して生成した標本ベクトルと、標本ベクトルの隠れマルコフモデルの状態遷移毎の出力確率とを、予め用意しておく。入力音声から入力ベクトルを生成し、この入力ベクトルに近い標本ベクトルを選出し、この標本ベクトルに対応する出力確率を読み出す。この出力確率に従って隠れマルコフモデルの出力確率を計算することで、入力ベクトルから状態遷移毎の出力確率を計算することを省略する。



(2)

特開平8-248985

1

【特許請求の範囲】

【請求項1】 状態遷移毎に出力確率が設定された隠れマルコフモデルを多数の標準音声の各々に対応させて形成し、入力音声に従って状態遷移毎の出力確率を順次計算して隠れマルコフモデルの出力確率を計算する音声認識装置において、

予め標準音声を時分割して生成した標準ベクトルが格納されたベクトル記憶手段を設け、予め状態遷移毎の出力確率が標準ベクトルに対応して格納された確率記憶手段を設け、入力音声を時分割して入力ベクトルを順次生成するベクトル生成手段を設け、生成される入力ベクトルに近い標準ベクトルを前記ベクトル記憶手段から選出するベクトル選出手段を設け、選出された標準ベクトルに対応する複数の出力確率を前記確率記憶手段から読み出す確率読出手段を設け、読み出された出力確率により隠れマルコフモデルの出力確率を計算する確率計算手段を設けたことを特徴とする音声認識装置。

【請求項2】 状態遷移毎に出力確率が設定された隠れマルコフモデルを多数の標準音声の各々に対応させて形成し、入力音声に従って状態遷移毎の出力確率を順次計算して隠れマルコフモデルの出力確率を計算する音声認識装置において、

予め標準音声を時分割して生成した標準ベクトルが格納されたベクトル記憶手段を設け、予め状態遷移毎の出力確率が標準ベクトルに対応して格納された確率記憶手段を設け、入力音声を時分割して入力ベクトルを順次生成するベクトル生成手段を設け、生成される入力ベクトルに近い標準ベクトルを前記ベクトル記憶手段から選出するベクトル選出手段を設け、選出された標準ベクトルに対応する複数の出力確率を前記確率記憶手段から読み出す確率読出手段を設け、読み出される複数の出力確率から上位候補を選出する確率選出手段を設け、上位候補として選出された出力確率に対して入力ベクトルから出力確率を再計算する確率再計算手段を設け、再計算された出力確率と読み出された他の出力確率とにより隠れマルコフモデルの出力確率を計算する確率計算手段を設けたことを特徴とする音声認識装置。

【請求項3】 ベクトル選出手段が標準ベクトルを選出する個数を複数に設定し、確率読出手段が読み出す複数の出力確率に対応する複数の標準ベクトルの重み付き内挿により統合する確率補正手段を設けたことを特徴とする請求項1又は2記載の音声認識装置。

【請求項4】 状態遷移毎に出力確率が設定された隠れマルコフモデルを多数の標準音声の各々に対応させて形成し、入力音声に従って状態遷移毎の出力確率を順次計算して隠れマルコフモデルの出力確率を計算する音声認識方法において、

標準音声を時分割して生成した標準ベクトルが格納されたベクトル記憶手段と、状態遷移毎の出力確率が標準ベクトルに対応して格納された確率記憶手段とを、予め形

2

成しておき、入力音声をベクトル生成手段により時分割して入力ベクトルを順次生成し、生成される入力ベクトルに近い標準ベクトルをベクトル選出手段により前記ベクトル記憶手段から選出し、選出された標準ベクトルに対応する複数の出力確率を確率読出手段により前記確率記憶手段から読み出し、読み出された出力確率により確率計算手段で隠れマルコフモデルの出力確率を計算するようにしたことを特徴とする音声認識方法。

【請求項5】 状態遷移毎に出力確率が設定された隠れマルコフモデルを多数の標準音声の各々に対応させて形成し、入力音声に従って状態遷移毎の出力確率を順次計算して隠れマルコフモデルの出力確率を計算する音声認識方法において、

標準音声を時分割して生成した標準ベクトルが格納されたベクトル記憶手段と、状態遷移毎の出力確率が標準ベクトルに対応して格納された確率記憶手段とを、予め形成しておき、入力音声をベクトル生成手段により時分割して入力ベクトルを順次生成し、生成される入力ベクトルに近い標準ベクトルをベクトル選出手段により前記ベクトル記憶手段から選出し、選出された標準ベクトルに対応する複数の出力確率を確率読出手段により前記確率記憶手段から読み出し、読み出される複数の出力確率から確率選出手段により上位候補を選出し、上位候補として選出された出力確率に対して確率再計算手段により入力ベクトルから出力確率を再計算し、再計算された出力確率と読み出された他の出力確率とにより確率計算手段で隠れマルコフモデルの出力確率を計算するようにしたことを特徴とする音声認識方法。

【請求項6】 ベクトル選出手段が複数の標準ベクトルを選出し、確率読出手段が読み出す複数の出力確率を確率補正手段が対応する複数の標準ベクトルの重み付き内挿により統合するようにしたことを特徴とする請求項4又は5記載の音声認識方法。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は、隠れマルコフモデルによる音声認識装置及び方法に関する。

【0002】

【従来の技術】 音声認識装置は、話者の発声が音声入力されると、この入力音声を単語や音節などとして認識する。このような音声認識装置には各種方式が提案されているが、その一つとして隠れマルコフモデルであるHMM(Hidden Markov Model)を利用した装置が提案されている。

【0003】 この場合、状態遷移毎に出力確率が設定された隠れマルコフモデルを、多数の標準音声の各々に予め設定しておく。入力音声に従って多数のHMMの出力確率を各々計算する。

【0004】 より具体的には、入力音声のHMMの出力確率を計算する場合、HMMには複数の状態が予め設定

50

(3)

特開平8-248985

されているので、これらの状態の遷移毎に出力確率を順次計算する。例えば、入力音声の時分割して入力ベクトル“x”が順次生成されるならば、この入力ベクトル“x”によりHMMの状態が“i”から“j”に遷移する*

* する場合、その出力確率“ $b_{ij}(x)$ ”は、以下の数式により計算される。

[0005]

[数1]

$$b_{ij}(x) = \sum_{a=1}^M \lambda_{ija} b_{ija}(x)$$

$$b_{ijn}(y) = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \exp \left\{ -\frac{1}{2} (y - \mu) \Sigma^{-1} (y - \mu)^T \right\}$$

【0006】上記数式において、“ λ_{ij} ”は分枝確率であり、“ λ_{ijn} ”はn番目の分布の出現確率である。

“ b_{ij} ”は出力確率密度であり、“ b_{ijn} ”はn番目の確率密度分布を示し、通常はガウス分布として仮定される。“ μ ”は平均ベクトルであり、標本音声をもとのフレームで時分割してからベクトル化した標本ベクトルに相当する。

【0007】

【発明が解決しようとする課題】上述のようにHMMの出力確率を計算することにより、入力音声を多数の標本音声の一つとして認識することができる。

【0008】しかし、このような音声認識では、HMMの出力確率の計算量は、HMMの個数とHMMの状態数と状態内の混合分布数とに比例するが、この計算量が多大であるため、処理速度の向上が困難である。特に、混合連続出力型HMMや、共分散行列に全共分散行列を利用する場合は、出力確率の計算量が膨大である。そこで、HMMの出力確率の計算量を削減するため、各種の方法が提案されている。

【0009】例えば、Bocchieriなどが提案した方法では、連続HMMで利用される混合ガウス分布の平均ベクトルを予めベクトル量子化しておき、入力音声から生成した入力ベクトルと量子化ベクトルとの距離情報を求める。この距離情報により混合ガウス分布のクラスタを予備選択し、このクラスタの選択に基づいて出力確率を計算している(Vector quantization for the efficient computation of continuous Density likelihoods, ICASSP-93, II, P692~P695, 1993, Bocchieri)。

【0010】また、渡辺などが提案した方法では、ガウス分布のクラスタの予備選択と、混合ガウス分布の出力確率の計算とに、木構造を導入している(木構造確率分布を用いた音声認識, 日本音響学会講演論文集, 1-8-7, 1993.10, 渡辺, 藤田, 高木, 山田, 服部, 敬)。

【0011】さらに、小森などが提案した方法では、認識に貢献度の高そうなHMMの状態を少数分布のHMMを利用して予備選択し、多数分布の出力確率を再計算する(少数分布HMMによる出力確率推定に基づいた効率的な混合連続分布HMM音声認識, 電子情報通信学会, SP94-52, 1994.10, 小森, 山田, 山本, 大洞)。

【0012】上述のような各種方法では、何れも実際に

計算する総分布数を削減しているが、未だ充分ではなく、さらに計算量を削減する方法が要望されている。

【0013】

【課題を解決するための手段】請求項1記載の発明は、状態遷移毎に出力確率が設定されたHMMを多数の標準音声の各々に対応させて形成し、入力音声に従って状態遷移毎の出力確率を順次計算してHMMの出力確率を計算する音声認識装置において、予め標本音声の時分割して生成した標本ベクトルが格納されたベクトル記憶手段を設け、予め状態遷移毎の出力確率が標本ベクトルに対応して格納された確率記憶手段を設け、入力音声の時分割して入力ベクトルを順次生成するベクトル生成手段を設け、生成される入力ベクトルに近い標本ベクトルを前記ベクトル記憶手段から選出するベクトル選出手段を設け、選出された標本ベクトルに対応する複数の出力確率を前記確率記憶手段から読み出す確率読出手段を設け、読み出された出力確率によりHMMの出力確率を計算する確率計算手段を設けた。

【0014】請求項2記載の発明は、状態遷移毎に出力確率が設定されたHMMを多数の標準音声の各々に対応させて形成し、入力音声に従って状態遷移毎の出力確率を順次計算してHMMの出力確率を計算する音声認識装置において、予め標本音声の時分割して生成した標本ベクトルが格納されたベクトル記憶手段を設け、予め状態遷移毎の出力確率が標本ベクトルに対応して格納された確率記憶手段を設け、入力音声の時分割して入力ベクトルを順次生成するベクトル生成手段を設け、生成される入力ベクトルに近い標本ベクトルを前記ベクトル記憶手段から選出するベクトル選出手段を設け、選出された標本ベクトルに対応する複数の出力確率を前記確率記憶手段から読み出す確率読出手段を設け、読み出される複数の出力確率から上位候補を選出する確率選出手段を設け、上位候補として選出された出力確率に対して入力ベクトルから出力確率を再計算する確率再計算手段を設け、再計算された出力確率と読み出された他の出力確率とによりHMMの出力確率を計算する確率計算手段を設けた。

【0015】請求項3記載の発明は、請求項1又は2記載の発明において、ベクトル選出手段が標本ベクトルを選出する個数を複数に設定し、確率読出手段が読み出す

10

20

30

40

50

【0023】前記データ処理部3には、RAM(Random Access Memory)などからなるデータ記憶部7が接続されており、このデータ記憶部7は、ベクトル記憶手段であるコードブック8と確率記憶手段である確率記憶部9とを有している。前記コードブック8は、前記コード選択部5に接続されており、前記確率記憶部9は、前記確率計算部6に接続されている。前記データ処理部3の確率計算部6には、ディスプレイなどからなる結果出力部10が接続されている。

(5)

特開平8-248985

7

【0024】本実施例の音声認識装置1は、詳細には後述するように、前記確率計算部6においてHMMの出力確率を計算するが、日本語の音節を標準音声として認識するため、“Left to Right”の連続HMMが、113音節毎に5状態により形成されている。

【0025】そして、前記コードブック8には、標準音声を所定フレームに時分割して生成した標本ベクトルとして、日本音響学界連続音声データベースの音声データからLBGアルゴリズムにより生成した256個のコードベクトル“ $y_1, \sim y_{256}$ ”が、識別子であるラベル“0”

20 “255”と共に予め格納されている。
【0026】前記確率記憶部9には、HMMの状態遷移毎の出力確率“ $b_{i,j}$ ”が、ラベル毎に格納されているので、これらの出力確率“ $b_{i,j}$ ”はラベルを介してコードベクトル“ y ”に対応している。コードベクトル“ y ”は256個、HMMは113個、分布(状態遷移)は4個であるので、図1に示すように、出力確率“ $b_{i,j}$ ”の総数は、115712個(=256×113×4)であり、1個のコードベクトル“ y ”に452個(=113×4)の出力確率“ $b_{i,j}$ ”が対応している。

【0027】このような構成において、本実施例の音声認識装置1では、話者が音声入力部2に対して日本語を発声すると、この入力音声はHMM法により日本語の音節として認識され、この認識結果が結果出力部10に出力される。このような音声認識装置1の音声認識方法を、図1及び図3に基づいて以下に順次詳述する。

【0028】まず、音声入力部2に音声が入力されると、ベクトル計算部4は、入力音声を所定フレームに時分割して入力ベクトル“ x_n ”を順次生成する。なお、この入力ベクトル“ x_n ”の“ n ”はフレームの番号である。

【0029】つぎに、コード選択部5は、コードブック8に格納された256個のコードベクトル“ $y_1, \sim y_{256}$ ”の全部と入力ベクトル“ x_n ”とのユークリッド距離“ $d_1, \sim d_{256}$ ”を計算し、このユークリッド距離“ d ”が最小の一つのコードベクトル“ y_r ”を選出する。このコードベクトル“ y_r ”は、入力ベクトル“ x_n ”に最も近い1個であるので、そのラベル“ r ”を確率計算部6に出力する。

【0030】そこで、確率計算部6は、ラベル“ r ”に対応する452個の出力確率“ $b_{i,j}$ ”を確率記憶部9から読み出し、この読み出された出力確率“ $b_{i,j}$ ”により113個のHMMの出力確率を計算することにより、このHMMの音節として入力音声は認識される。このような音節の認識が連続的に実行されることにより、連続HMMにより入力音声は認識され、これが認識結果として結果出力部10に出力される。

【0031】つまり、本実施例の音声認識装置1の音声認識方法では、HMMの状態遷移毎の出力確率を予め計算してコードベクトル毎に確率記憶部9に格納してお

8

き、入力音声に最も近いコードベクトルに対応して確率記憶部9から出力確率を読み出し、この読み出した出力確率によりHMMの出力確率を計算する。このため、HMMの出力確率を計算する際、状態遷移毎の出力確率を入力ベクトルから計算する必要がないので、全体の計算量が極めて少なく処理速度が高速である。

【0032】なお、本実施例の音声認識装置1の音声認識方法では、上述のように確率記憶部9から読み出した出力確率によりHMMの出力確率を計算することを例示したが、本発明は上記実施例に限定されるものではない。

【0033】例えば、確率計算部6に、確率選出手段と確率再計算手段と確率計算手段とを實現し、図4に示すように、確率記憶部9から読み出される452個の出力確率から上位160個を確率選出手段により上位候補として選出する。この上位候補として選出された出力確率に対しては、確率再計算手段により従来と同様に入力ベクトルから出力確率を再計算し、この再計算された出力確率と読み出された他の出力確率とにより、確率計算手段でHMMの出力確率を計算する。なお、上位候補として選出されない292個の出力確率は、そのままHMMの出力確率を計算する。

【0034】この場合は、入力音声と一致する確率が高いHMMのみ、入力ベクトルから出力確率を計算するので、全体の計算量が少なく、認識精度が良好である。なお、ここでは入力音声と一致する確率が低いHMMに対しては、確率記憶部9から読み出した出力確率によりHMMの出力確率を計算することにより、計算量を低く維持しながらも、認識精度を向上させているが、例えば、このようなHMMに対する計算は中止することにより、より計算量を削減することも可能である。

【0035】さらに、本実施例の音声認識装置1の音声認識方法では、コードブック8に格納された256個のコードベクトル“ $y_1, \sim y_{256}$ ”から入力ベクトル“ x_n ”に最も近い1個を選出し、この1個のコードベクトル“ y_r ”に対応した452個の出力確率“ $b_{i,j}$ ”を確率記憶部9から読み出すことを例示したが、本発明は上記実施例に限定されるものでもない。

【0036】例えば、コード選択部5がコードベクトル“ y_r, y_s ”を選出する個数を2個に設定しておき、確率計算部6が読み出す“2×452”個の出力確率“ $b_{i,j}(y_r), b_{i,j}(y_s)$ ”を、2個のコードベクトル“ y_r, y_s ”の重み付き内挿により、確率補正手段で“ $b_{i,j}(y)$ ”として統合することも可能である。つまり、入力ベクトル“ x_n ”とコードベクトル“ y_r ”とのユークリッド距離が“ d_r ”、入力ベクトル“ x_n ”とコードベクトル“ y_s ”とのユークリッド距離が“ d_s ”ならば、これらのユークリッド距離の比率を重みとして“2×452”個の出力確率“ $b_{i,j}(y_r), b_{i,j}(y_s)$ ”を内挿することにより、下記の数式のように、“452”

50

9

個の出力確率 $b_{ij}(y)$ を算出することができる。

【0037】

【数2】

$$b_{ij}(y) = \frac{d_{y_1} b_{ij}(y_1) + d_{y_2} b_{ij}(y_2)}{d_{y_1} + d_{y_2}}$$

【0038】この場合、入力ベクトル“x”に一致するコードベクトル“y”がコードブック8に存在しなくても、入力ベクトル“x”に一致するコードベクトル“y”を擬似的に出力することになる。そして、このコードベクトル“y”の出力確率 $b_{ij}(y)$ も重み付き内挿により擬似的に出力できるので、コードベクトル“y”が少数でも良好な音声認識を実現することができる。

【0039】また、本実施例の音声認識装置1では、ベクトル記憶手段をコードブック8とし、ここに格納する標本ベクトルをコードベクトルとすることを例示したが、本発明は上記実施例に限定されるものでもなく、HMMの出力確率分布の平均ベクトルを標本ベクトルとすることも可能である。この場合、詳細には後述するよう

に、認識精度が低下する場合があるが、標本ベクトルを*

サンプリング周波数 12 (kHz)

フレームシフト 8.7 (ms)

ハミング窓長 23.22 (ms)

プリエンファシス 1-z⁻¹

分析 14次LPC分析

特徴パラメータ 10次LPCメルケプストラム係数

として設定した。

【0043】そして、このような音声認識装置1で認識する入力音声を、東北大-松下技研の単語音声データベースの男性話者15名の各205単語とした。まず、入力音声に対応してHMMの状態遷移毎の出力確率を全部計算する従来方式では、認識率が92.6(%)となった。

【0044】一方、入力音声に対応してコードブックから読み出したベクトルによりHMMを計算する音声認識方法において、256個のコードベクトルを格納した第一のコードブックを使用したところ、その認識率は85.7(%)となった。同様に、512個のコードベクトルを格納した第二のコードブックを使用した場合の認識率は87.0(%)、452個の平均ベクトルを格納した第三のコードブックを使用した場合の認識率は53.9(%)となった。

【0045】つぎに、コードブックから読み出したベクトルから160個の上位候補を選出し、この上位候補のみ出力確率を入力ベクトルから再計算する音声認識方法を実行した。すると、第一のコードブックを使用した場合の認識率は91.8(%)、第二のコードブックを使用した場合の認識率は92.3(%)、第三のコードブックを使用した場合の認識率は90.8(%)となった。この場合、処

(6)

特開平8-248985

10

*新規に作成する必要がないので、音声認識装置1の実現が容易となる。

【0040】ここで、本実施例の音声認識装置1の試作品による実験結果を以下に説明する。まず、日本語の13音節に対し、HMMを“Left to Right”の連続出力分布型として5状態4分布に設定した。このHMMは、全共分散行列を用いた10次元の正規分布(混合分布数は1)で出力確率分布を表す。

【0041】また、コードブック8としては、三種類を用意した。まず、第一のコードブックには、日本音響学界連続音声データベースの30名の男性話者の音声データから、LBGアルゴリズムにより生成した256個のコードベクトルを格納した。この時、合計で約100万フレームを使用した。第二のコードブックには、同様にして512個のコードベクトルを格納し、第三のコードブックには、HMMの出力確率分布の452個の平均ベクトルを格納した。なお、このようなものは一般的にはコードブックと呼称しないが、ここでは簡単のためにコードブックと呼称する。

【0042】さらに、この音声認識装置1の分析条件を、

理時間は従来方式が1秒強であるのに対し、本素方式では0.5秒強となった。

【0046】つまり、本素の音声認識装置1の音声認識方法では、従来方式と同等の認識率を実現しながらも、従来方式より極めて短時間に処理を完了できることが、実証された。なお、図5及び図6に示すように、再計算個数を増加させると、本素方式の認識率は従来方式に漸近的に近接するが、処理時間は再計算個数に略線形に比例するので、再計算個数は必要最小限とすることが好ましい。

【0047】また、このように再計算個数を“0”から順次増加させると、認識率は一度低下してから上昇することが判明した。この原因としては、コードベクトルを生ベクトルの全体の平均で計算したために特徴を損失したとことと、ベクトル量子化による誤差が大きいためにと考えられる。このようなことを解消するため、生ベクトルから512個の代表ベクトルを選出し、これをコードブックに設定して音声認識を実験したところ、図5に示すように、認識率の向上が認められた。

【0048】さらに、代表ベクトルを256個とし、入力ベクトルに近い2個の代表ベクトルを選出し、これら2個の代表ベクトルの重み付き内挿により出力確率を統合

(7)

特開平8-248985

11

することも実験した。すると、認識率は再計算が70個の場合で 90.8(%)、再計算が120個の場合で 92.3(%) となり、重み付き内挿により標本ベクトルが少数でも良好な音声認識が実現できることが実証された。

【0049】

【発明の効果】請求項1及び4記載の発明は、標本音声の時分割して生成した標本ベクトルが格納されたベクトル記憶手段と、状態遷移毎の出力確率が標本ベクトルに対応して格納された確率記憶手段とを、予め形成しておき、入力音声ベクトル生成手段により時分割して入力ベクトルを順次生成し、生成される入力ベクトルに近い標本ベクトルをベクトル選出手段によりベクトル記憶手段から選出し、選出された標本ベクトルに対応する複数の出力確率を確率読出手段により確率記憶手段から読み出し、読み出された出力確率により確率計算手段によりHMMの出力確率を計算することにより、HMMの出力確率を計算する際、入力ベクトルから出力確率を計算しないので、全体の計算量を削減することができ、音声認識の処理速度を向上させることができる。

【0050】請求項2及び5記載の発明は、標本音声の時分割して生成した標本ベクトルが格納されたベクトル記憶手段と、状態遷移毎の出力確率が標本ベクトルに対応して格納された確率記憶手段とを、予め形成しておき、入力音声ベクトル生成手段により時分割して入力ベクトルを順次生成し、生成される入力ベクトルに近い標本ベクトルをベクトル選出手段によりベクトル記憶手段から選出し、選出された標本ベクトルに対応する複数の出力確率を確率読出手段により確率記憶手段から読み出し、読み出される複数の出力確率から確率選出手段により上位候補を選出し、上位候補として選出された出力確率に対して確率再計算手段により入力ベクトルから出力確率を再計算し、再計算された出力確率と読み出され*

12

*た他の出力確率とにより確率計算手段によりHMMの出力確率を計算することにより、このようにHMMの出力確率を計算する際、一部の入力ベクトルのみから出力確率を計算するので、全体の計算量を削減することができ、音声認識の処理速度を向上させることができる。

【0051】請求項3及び6記載の発明は、ベクトル選出手段が複数の標本ベクトルを選出し、確率読出手段が読み出す複数の出力確率を確率補正手段が対応する複数の標本ベクトルの重み付き内挿により統合することにより、入力ベクトルに一致する標本ベクトルを仮想的に出力することができ、この標本ベクトルの出力確率も重み付き内挿により仮想的に出力することができるので、標本ベクトルが少数でも入力音声の認識精度が良好である。

【図面の簡単な説明】

【図1】本発明の一実施例の音声認識装置の音声認識方法を示す模式図である。

【図2】音声認識装置を示すブロック図である。

【図3】第一の音声認識方法を示すフローチャートである。

【図4】第二の音声認識方法を示すフローチャートである。

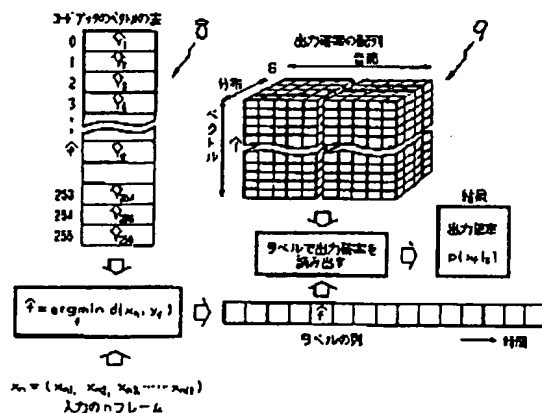
【図5】音声認識装置の試作品における再計算回数と認識率との関係を示す特性図である。

【図6】音声認識装置の試作品における再計算回数と処理時間との関係を示す特性図である。

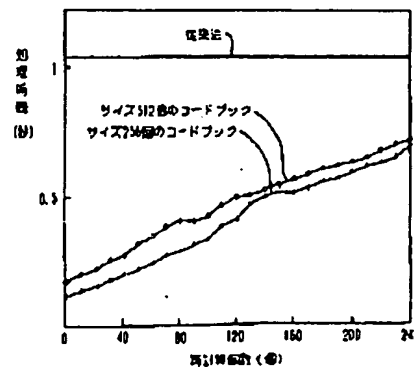
【符号の説明】

- 1 音声認識装置
- 4 ベクトル生成手段
- 5 ベクトル選出手段
- 8 ベクトル記憶手段
- 9 確率記憶手段

【図1】



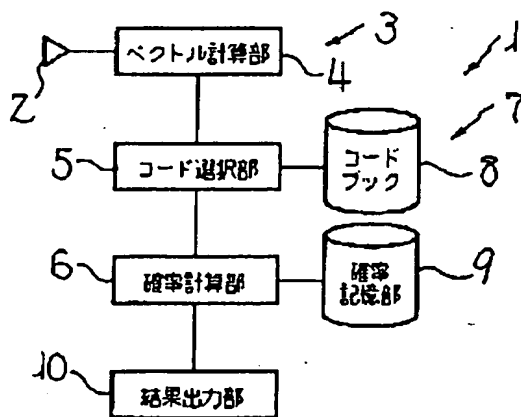
【図6】



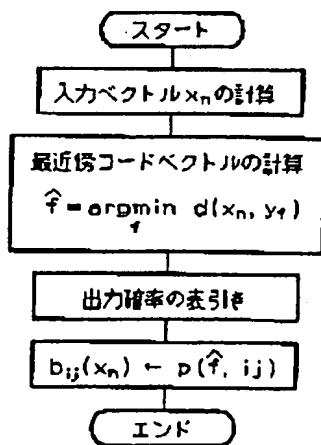
(8)

特開平8-248985

【図2】



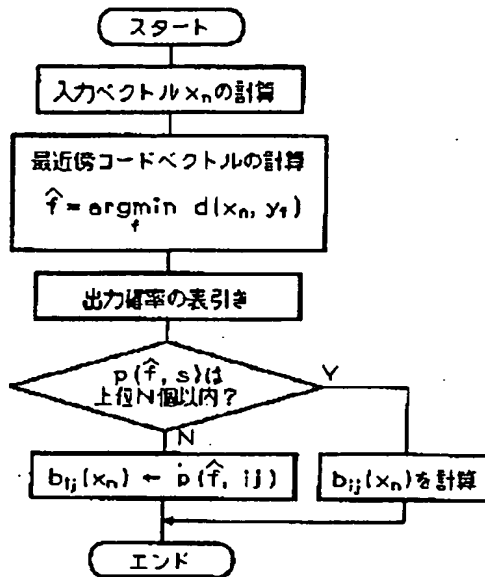
【図3】



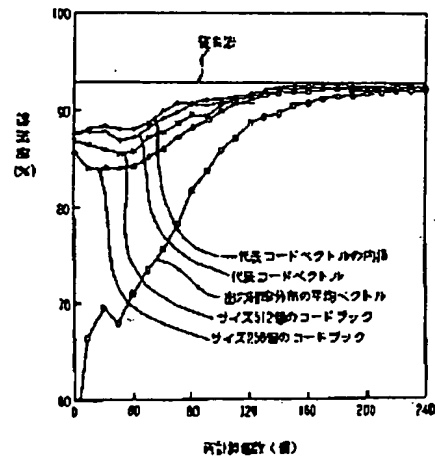
(9)

特開平8-248985

【図4】



【図5】



**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☒ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.